Energy-conserving finite difference schemes for nonlinear strings

1

Energy-conserving Finite Difference Schemes for Nonlinear Strings

Stefan Bilbao* and Julius O. Smith III

Abstract

In this paper, we discuss numerical methods for the simulation of nonlinear string vibration. In a musical setting, such vibrations occur under large-amplitude conditions. Though it is simple enough to write down a nonlinear equation for the motion of such a string, the construction of numerical methods, such as finite difference schemes, becomes a much more delicate problem, in particular because of the number of possible discretization approaches, and the lack of spectral analysis tools which are often used to determine stability bounds in the linear case. For this reason, we turn towards more general techniques, which are based around energetic principles, sometimes known collectively as the energy method. Indeed, it is possible to transfer the complete energetic behavior of the nonlinear string to discrete time in such a way that a discrete analogue of the energy is preserved. This leads, in turn, to an energy-based stability guarantee. Such techniques rely in no way on spectral analysis. Various other issues, such as parasitic oscillations, boundary termination, loss modelling and implicit difference schemes are also discussed. Simulations are presented.

Keywords

string vibration, finite difference schemes, energy method, sound synthesis, Kirchhoff-Carrier equation

I. Introduction

The usual starting point for sound synthesis methods for stringed instruments is a numerical treatment of the 1D wave equation. This model equation, though simplified, captures many of the salient features of transverse wave propagation on a string, and is very nearly correct when amplitudes are small; it may be complemented by various additional terms which model loss and dispersion, which are effects of great perceptual relevance. Still, it is a linear model, and falls short when it is desired to simulate or synthesize sounds which result from large-amplitude vibration.

In this paper, we discuss a nonlinear generalization of the wave equation, sometimes called the Kirchhoff-Carrier equation, which is perhaps the simplest possible nonlinear model for string vibration; in essence, the tension in the string is "modulated" due to time variations in the length of the string. These changes in tension lead to changes in wave speed, which ultimately influence the perceived pitch of the waveform produced. This model has recently been used as the point of departure for synthesis models based on digital waveguides [1], efficient structures which have long been used for synthesis purposes for linear strings; wave propagation in a tension-modulated string is treated as linear along the length of the string, with nonlinear effects modeled via state-dependent digital filters inserted at the string terminations, which serve to modulate the effective string length [2], [3], [4].

Another simulation approach, based on finite difference schemes [5], [6], [7], also has been used extensively for string synthesis; the earliest work was that of Ruiz [8] and Hiller and Ruiz [9], followed by that of Bacon and Bowsher [10], and later the comprehensive work of Chaigne [11] and Chaigne and Askenfelt [12] and subsequent authors [13]. In all these cases, the model system is the linear 1D wave equation, with certain perturbation terms added. Analysis of the resulting finite difference schemes has been carried out through the use of Fourier-based or von Neumann analysis [7], [5], and as such, is only of use for linear schemes. There do, however, exist families of analysis techniques which may be used for stability analysis, without any reference to the frequency domain. Such techniques involve the identification of so-called conserved quantities in a difference scheme (and in particular, energy), which may carry over from the model system itself; such conserved quantities can lead to numerical stability bounds, even in the nonlinear case. Such techniques are often collectively referred to as the *energy method* [6], [14]; a particularly useful tutorial reference, for the case of the nonlinear Klein Gordon equation (which is also a nonlinear generalization of the wave equation) is given in [15]. Though these methods have attained widespread currency throughout computational physics, they have not as yet been used for the analysis of numerical methods for musical systems. (We do note, however, that energetic ideas have been applied to a numerical method for a kettledrum [16], though in a very different sense from that presented here, and to developing energy-normalized digital filter structures [17], [1].) The determination of global stability conditions for digital sound synthesis is a critical concern, especially if such a digital model is to be used in a live performance setting; frequency-domain methods simply do not allow a satisfactory guarantee of stability in the nonlinear case. It is also important to note that while the extended digital waveguide approach noted at the end of the above paragraph does indeed yield an extremely efficient and stable nonlinear synthesis method, it does not, strictly speaking, give a solution to the Kirchhoff-Carrier equation. We have decided here to stick as closely as possible to the model equation itself, thus sacrificing the efficiency a waveguide type model permits. Further comments on the distinction between waveguides

S. Bilbao (email: s.bilbao@qub.ac.uk) is with the Sonic Arts Research Centre, the Queen's University, Belfast, Northern Ireland, and J. O. Smith III is with the Center for Computer Research in Music and Acoustics, Department of Music, Stanford University, Stanford, CA, 94305-8180.

and finite difference schemes in the present case of nonlinear strings are provided in the concluding remarks. Some preliminary work on this topic was presented in [18].

In Section II, we introduce the 1D wave equation, as well as the Kirchhoff-Carrier equation, and reduce both to first-order systems. We then examine energy conservation properties of both equations. In Section III, after a cursory review of grid functions and finite difference operators, we present finite difference schemes for both the wave equation and the Kirchhoff-Carrier equation. We then proceed to derive conserved quantities for both schemes, as well as the conditions under which they are positive (and hence have the interpretation of energies). This positivity condition then leads to bounds on the size of the state of the difference scheme, which serves as a numerical stability guarantee. We then spend some time looking at the problem of spurious high-frequency oscillations and the conditions under which they are present, and conclude the section with a treatment of fixed boundary conditions. Numerical results are presented in Section IV. Other auxiliary directions are pursued in the Appendix, first the modelling of loss, both in the model equation and in the resulting difference scheme, then an alternative implicit scheme, for which a conserved energetic quantity may also be derived.

II. The Wave Equation and the Kirchhoff-Carrier Equation

For a string undergoing transverse motion in a single polarization, the 1D wave equation is a first approximation. It can be written simply as

$$\rho \frac{\partial^2 u}{\partial t^2} = T_0 \frac{\partial^2 u}{\partial x^2} \tag{1}$$

Here, $t \geq 0$ is a time variable, $x \in [0, L]$ is a space variable, and $u(x,t)$ is the transverse string displacement. ρ is the linear mass density of the string, and T_0 is the tension applied to the string; both are assumed constant. Given two initial conditions, $u(x,0)$ and $\frac{\partial u}{\partial t}(x,0)$, and appropriate boundary conditions, the solution to the wave equation exists and is unique for all future times $t > 0$. The wave equation can be complemented by additional linear terms which model loss and dispersion[12], [13].

A nonlinear generalization of the wave equation (1) to model large-amplitude vibration in stiff strings was first put forth by Kirchhoff in 1883 [19], and subsequently rediscovered by Carrier [20]. It can be written as

$$\rho \frac{\partial^2 u}{\partial t^2} = \left(T_0 + \frac{EA}{2L} \int_0^L \left(\frac{\partial u}{\partial x} \right)^2 dx \right) \frac{\partial^2 u}{\partial x^2} \tag{2}$$

where E is Young's modulus, and A is the string cross-sectional area [3],[21], [22]. (We note that many slightly different forms of these equations have been proposed; see, e.g., [23] and [24] for some examples.) Essentially, the tension in the string consists of the applied tension T_0 plus an additional contribution due to the significant change in length of the string under large-amplitude conditions. The Kirchhoff-Carrier equation can be shown to be the result of various simplifying assumptions applied to a more complete model of string vibration [25], [26] which takes into account the two transverse polarizations as well as longitudinal motion and the coupling among them. The most important assumption here, often referred to as *Anand's argument* [27] is that longitudinal wave speed is significantly greater than transverse speed; this is justified for most cases of musical interest.

As it turns out, some (but not all) of what can be said about the behavior of the solutions of (1) can be generalized to deal the solutions of (2). This is also true of certain numerical methods, and we will present a parallel treatment of the numerical solution of these equations in this paper.

A. First-order Systems

For analysis purposes, and for the development of numerical methods (such as finite difference schemes) [5], it is useful to rewrite the wave equation (1) as a first-order system, i.e.,

$$\frac{\partial p}{\partial t} = c_0 \frac{\partial q}{\partial x} \tag{3a}$$

$$\frac{\partial q}{\partial t} = c_0 \frac{\partial p}{\partial x} \tag{3b}$$

where we have introduced the new variables p and q, as well as the *wave speed* c_0, all defined by

$$p = \sqrt{\rho} \frac{\partial u}{\partial t} \qquad q = \sqrt{T_0} \frac{\partial u}{\partial x} \qquad c_0 \triangleq \sqrt{\frac{T_0}{\rho}}$$

Both p and q have units proportional to root energy density. This system is of the form of the transmission line, or telegrapher's equations [28].

Like the wave equation, (2) can be written as a first order system

$$\frac{\partial p}{\partial t} = c_0 G \frac{\partial q}{\partial x} \tag{4a}$$

$$\frac{\partial q}{\partial t} = c_0 \frac{\partial p}{\partial x} \tag{4b}$$

where the variables p and q are defined as before, and where we have introduced the quantity

$$G \triangleq \left(1 + B \int_0^L q^2 dx\right) \tag{4c}$$

with $B \triangleq EA/2LT_0^2$. System (4) is thus identical to (3), aside from the factor G. We also note that G, as defined, is a function of time alone, through its dependence on q.

B. Energetic Analysis

Since the wave equation, (or the equivalent first-order system) is linear and shift-invariant, it is possible to perform a complete analysis and arrive at solutions using frequency domain techniques, such as Laplace and Fourier transforms [5], [6], [7]. As we would like to be able to generalize the analysis to the nonlinear case, we will make use of non-spectral techniques, and in particular, the energy method [6][1].

As the string, as presented here, in either the linear or nonlinear case, is defined over the interval $x \in [0, L]$, it is important to specify boundary conditions. For preliminary energetic analysis, we assume periodic boundary conditions, i.e., $p(0,t) = p(L,t)$ and $q(0,t) = q(L,t)$; in other words, the string is a boundaryless "loop." We will return to the case of fixed boundary termination in Section III-F.

In the linear case, one can perform an energetic analysis as follows: multiply (3a) by p to get

$$p\frac{\partial p}{\partial t} = c_0 p \frac{\partial q}{\partial x} \tag{5}$$

Integrating over the range $x \in [0, L]$ gives

$$\frac{d}{dt}\int_0^L \frac{1}{2}p^2 dx = c_0 \int_0^L p\frac{\partial q}{\partial x} dx = -c_0 \int_0^L q\frac{\partial p}{\partial x} dx + c_0 pq|_0^L = -\frac{d}{dt}\int_0^L \frac{1}{2}q^2 dx \tag{6}$$

where, in the second equality above, we have used integration by parts, and in the third equality, the defining equation (3b), as well as periodicity of boundary conditions. Note that the above would also remain true for boundary conditions $p = 0$ or $q = 0$ (i.e., fixed, or free) at either end of the string.

Defining now the function $E(t)$ by

$$E(t) = \frac{1}{2}\|p\|^2 + \frac{1}{2}\|q\|^2 \tag{7}$$

where we have used the notation $\|f\| = \left(\int_0^L f^2 dx\right)^{1/2}$ for square-integrable functions $f \in L^2(0, L)$, we clearly have that

$$\frac{dE}{dt} = 0 \qquad \Rightarrow \qquad E(t) = E(0)$$

Here, $E(t)$ has the interpretation of the total string energy; the term $\|p\|^2/2$ represents the kinetic energy, and the term $\|q\|^2/2$ the potential energy. As the wave equation does not include a loss mechanism, this energy must remain constant (and equal to the initial energy present in the system); we examine a simple loss mechanism in Appendix A. In a numerical setting, it is useful to have such an energy conservation property (i.e., a positive definite function of the state), as it can be used to ensure numerical stability, as we will show shortly. To this end, we note that as $E(t)$ is clearly non-negative, we can immediately bound the size of p and q in an L^2 norm by

$$\|p\|, \|q\| \leq \sqrt{2E(t)} = \sqrt{2E(0)}$$

[1]In the case of the Kirchhoff-Carrier equation, it is in fact possible to perform a "half-way" spectral analysis in the spatial dimension, due to the lack of spatial dependence of the nonlinearity. We would like here to emphasize the use of energetic analysis of PDEs and difference schemes, especially considering that more realistic models of string vibration are not amenable to any type of spectral analysis. For this reason we avoid the use of transformed quantities whenever possible. An exception is the analysis of oscillatory behavior, to be presented in Section III-E, for which spectral tools offer a convenient approach.

The analysis of the Kirchhoff-Carrier system (4) is very similar; after multiplying (4a) by p and integrating over $x \in [0, L]$, we may proceed again through a set of equalities[2],

$$\frac{d}{dt}\int_0^L \frac{1}{2}p^2 dx = c_0 G \int_0^L p\frac{\partial q}{\partial x}dx = c_0 G\left(-\int_0^L q\frac{\partial p}{\partial x}dx + pq|_0^L\right) = -G\frac{d}{dt}\int_0^L \frac{1}{2}q^2 dx \tag{8}$$

Noting, from (4c), that $G = 1 + B\|q\|^2$, we thus have

$$\frac{d}{dt}\left(\frac{1}{2}\|p\|^2\right) + \left(1 + B\|q\|^2\right)\frac{d}{dt}\left(\frac{1}{2}\|q\|^2\right) = 0 \qquad \Rightarrow \qquad \frac{dE_{KC}}{dt} = 0 \tag{9}$$

where E_{KC} is defined by

$$E_{KC} = \frac{1}{2}\|p\|^2 + \frac{1}{2}\left(1 + \frac{B}{2}\|q\|^2\right)\|q\|^2 \tag{10}$$

E_{KC}, like E, is positive for any choice of p and q, and remains constant as time progresses. It also may be used to bound the sizes of p and q, by

$$\|p\| \leq \sqrt{2E_{KC}} \qquad \|q\| \leq \sqrt{\frac{-1 + \sqrt{1 + 4BE_{KC}}}{B}} \tag{11}$$

In the next section, we show how this continuous time and space analysis may be transferred to discrete time; the bounds on the size of the state in terms of a positive conserved energy become extremely useful as a guarantee of numerical stability.

III. Finite Difference Schemes

Before proceeding to a difference scheme for the transmission line equations, we provide here some basic facts about grid functions and difference operators [6].

A. Grid Functions and Difference Operators

A real-valued grid function f_i^n, employed in a 1D finite difference scheme, is to be viewed as an approximation to a continuous time/space variable $f(x_i, t^n)$, at the coordinates $x_i = ih$, $t^n = nk$, for integer i and n; here h is the grid spacing, and k is the time-step (both assumed constant). We assume f_i^n to be limited to a domain consisting of N points, e.g., $i = 0, \ldots, N-1$. Because we will be dealing primarily with difference schemes of the *interleaved* variety, it is also to helpful to define grid functions such as $g_{i+1/2}^{n+1/2}$, again for integer n and $i = 0, \ldots, N-1$. Such a grid function will represent an approximation to a continuous function at time $t^{n+\frac{1}{2}} = (n + 1/2)k$ and at location $x_{i+\frac{1}{2}} = (i + 1/2)h$.

The forward time difference operator δ_{t+} and forward time-averaging operator μ_{t+} are defined by

$$\delta_{t+} f_i^n = \frac{1}{k}(f_i^{n+1} - f_i^n) \qquad \mu_{t+} f_i^n = \frac{1}{2}(f_i^{n+1} + f_i^n)$$

The identities

$$(\mu_{t+} f_i^n)(\delta_{t+} f_i^n) = \delta_{t+}\left(\frac{1}{2}(f_i^n)^2\right) \tag{12}$$

$$\mu_{t+}\delta_{t+} f_i^n = \frac{1}{2k}(f_i^{n+2} - f_i^n) \tag{13}$$

follow immediately from the definitions above.

Forward and backward spatial difference operators δ_{x+} and δ_{x-} are defined by

$$\delta_{x+} f_i^n = \frac{1}{h}(f_{i+1}^n - f_i^n) \qquad \delta_{x-} f_i^n = \frac{1}{h}(f_i^n - f_{i-1}^n)$$

For periodic boundary conditions, the spatial indices of the grid function are to be taken modulo N. For instance, $\delta_{x+} f_{N-1}^n = \frac{1}{h}\left(f_0^n - f_{N-1}^n\right)$. We also note that the operators δ_{t+}, μ_{t+}, δ_{x+}, and δ_{x-} all commute.

[2]We remind the reader that the quantity G, as defined in (4c), has no spatial dependence, and can be thought of as a function of the entire string state at a given time instant. This lack of spatial dependence of the nonlinearity is of crucial importance in the energetic analysis in this section, and also is the basis of the efficiency of the extended digital waveguide implementation [2].

For the subsequent energetic analysis of difference schemes, it is useful to define an inner product at time step n between two real-valued grid functions f_i^n and g_i^n by

$$\langle f^n, g^n \rangle = \sum_{i=0}^{N-1} h f_i^n g_i^n$$

(The grid function g_i^n above may be replaced by a grid function $g_{i+1/2}^{n+1/2}$, interleaved with respect to f_i^n without affecting the above definition.) An l^2 norm, for square-summable sequences, then follows as

$$\|f^n\| = \langle f^n, f^n \rangle^{1/2}$$

We also recall the Cauchy-Schwartz inequality,

$$|\langle f^n, g^n \rangle| \leq \|f^n\| \|g^n\| \tag{14}$$

as well as the triangle inequality, $\|f^n + g^n\| \leq \|f^n\| + \|g^n\|$, which implies, in particular, that

$$\|\delta_{x-} f^n\| \leq \frac{2}{h} \|f^n\| \tag{15}$$

The useful identity

$$\langle f^n, \delta_{x-} g^n \rangle = -\langle \delta_{x+} f^n, g^n \rangle \tag{16}$$

is the discrete analogue of integration by parts; this definition holds for periodic boundary conditions.

B. Interleaved Finite Difference Schemes

The transmission-line equations (3) are most easily approached, numerically, through simple centered, explicit and interleaved finite difference schemes of the Yee variety [29], [30]. The variables p and q are approximated at alternating spatial locations and time-steps. We thus define grid functions p_i^n and $q_{i+\frac{1}{2}}^{n+\frac{1}{2}}$, both for integer n and $i = 0, \ldots, N-1$. p is calculated at coordinates x and t at even multiples of $h/2$ and $k/2$, respectively, and q at odd multiples.

A difference scheme corresponding to (3) is then

$$\delta_{t+} p_i^{n-1} = c_0 \delta_{x-} q_{i+\frac{1}{2}}^{n-\frac{1}{2}} \tag{17a}$$

$$\delta_{t+} q_{i+\frac{1}{2}}^{n-\frac{1}{2}} = c_0 \delta_{x+} p_i^n \tag{17b}$$

This scheme is consistent with system (3), and second order accurate, by virtue of centering of difference approximations.

System (4) can also be approximated using an interleaved difference scheme like (17), but now of the form

$$\delta_{t+} p_i^{n-1} = c_0 \mathcal{G}^{n-\frac{1}{2}} \delta_{x-} q_{i+\frac{1}{2}}^{n-\frac{1}{2}} \tag{18a}$$

$$\delta_{t+} q_{i+\frac{1}{2}}^{n-\frac{1}{2}} = c_0 \delta_{x+} p_i^n \tag{18b}$$

Here, the function $\mathcal{G}^{n-\frac{1}{2}}$ is some second-order accurate approximation to the continuous time function G as given in (4c); notice, in particular, that it is not a grid function, as it depends only on time. We leave its exact form unspecified for the moment (one choice is given in (25)), and will return to it shortly.

Though we will not discuss initialization in this paper, it is simple enough to see that either of the two systems (3) or (4) requires the specification of two initial conditions, $p(x,0)$ and $q(x,0)$. The difference schemes above also require two conditions, p_i^0, and $q_{i+\frac{1}{2}}^{\frac{1}{2}}$. The first of these can be simply set to be a sampled version of $p(x,0)$, but the second condition requires more care, as it represents an approximation to q at time $k/2$; we may either set it to a sampled version of $q(x,0)$, and accept the first-order error that is incurred, or we may develop a specialized routine, to be used only at the initial time step (making use of one-sided rather than centered differences).

C. Conserved Quantities

Energetic analysis of the system (17) can be carried out in a near exact analogy with that of the continuous time and space system. First, multiply (17a) by $\mu_{t+} p_i^{n-1}$, to get

$$(\mu_{t+} p_i^{n-1})(\delta_{t+} p_i^{n-1}) = c_0 (\mu_{t+} p_i^{n-1})(\delta_{x-} q_{i+\frac{1}{2}}^{n-\frac{1}{2}}) \tag{19}$$

and by applying identity (12), we arrive at

$$\frac{1}{2}\delta_{t+}(p_i^{n-1})^2 = c_0(\mu_{t+}p_i^{n-1})(\delta_{x-}q_{i+\frac{1}{2}}^{n-\frac{1}{2}}) \tag{20}$$

We then sum over i and multiply by h, to get

$$\frac{1}{2}\delta_{t+}\|p^{n-1}\|^2 = c_0\langle\mu_{t+}p^{n-1}, \delta_{x-}q^{n-\frac{1}{2}}\rangle \tag{21}$$

and, continuing,

$$\begin{aligned}\frac{1}{2}\delta_{t+}\|p^{n-1}\|^2 &= -c_0\langle\delta_{x+}\mu_{t+}p^{n-1}, q^{n-\frac{1}{2}}\rangle \\ &= -c_0\langle\mu_{t+}\delta_{x+}p^{n-1}, q^{n-\frac{1}{2}}\rangle \\ &= -\langle\mu_{t+}\delta_{t+}q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle \\ &= -\frac{1}{2k}\langle q^{n+\frac{1}{2}} - q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle = -\frac{1}{2}\delta_{t+}\langle q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle\end{aligned} \tag{22}$$

where the above steps follow from identity (16), commutativity of the operators μ_{t+} and δ_{x+}, the defining equation (17b), and identity (13) respectively. It is easy to conclude that

$$\delta_t\left(\frac{1}{2}\|p^{n-1}\|^2 + \frac{1}{2}\langle q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle\right) = \delta_t\left(\frac{1}{2}\|p^{n-1}\|^2 + \frac{1}{2}\|q^{n-\frac{1}{2}}\|^2 - \frac{\lambda h}{2}\langle\delta_{x+}p^{n-1}, q^{n-\frac{1}{2}}\rangle\right) = 0 \tag{23}$$

where we have introduced the important parameter

$$\lambda \triangleq \frac{kc_0}{h}$$

which is sometimes called the *Courant number*. Thus the quantity $\mathcal{E}^n$, defined by

$$\mathcal{E}^n = \frac{1}{2}\left(\|p^n\|^2 + \|q^{n+\frac{1}{2}}\|^2 - \lambda h\langle\delta_{x+}p^n, q^{n+\frac{1}{2}}\rangle\right) \tag{24}$$

which is dependent only on the current state $p^n, q^{n+\frac{1}{2}}$ of the scheme, must remain constant as the recursion progresses. It is thus a *conserved quantity* of difference scheme (17). $\mathcal{E}^n$ is very nearly the discrete-time equivalent of the continuous-time energy E defined in (7), but we have yet to show the conditions under which it is positive; we return to this analysis in the next section.

The analysis of difference system (18) for the Kirchhoff-Carrier equation is identical, except for the extra factor of $\mathcal{G}^{n-\frac{1}{2}}$, which is carried through up until the form given in (22). Now, however, we have

$$\frac{1}{2}\delta_{t+}\|p^{n-1}\|^2 = -\frac{1}{2}\mathcal{G}^{n-\frac{1}{2}}\delta_{t+}\langle q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle$$

and cannot proceed further to a conservation law until the form of $\mathcal{G}^{n-\frac{1}{2}}$ is specified; whatever choice we make must be consistent with the definition of G from (4c), and, furthermore, should be second order accurate in order that difference scheme (18) remain second order accurate overall. One possible choice for such an approximation is

$$\mathcal{G}^{n-\frac{1}{2}} = 1 + B\mu_{t+}\langle q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle \tag{25}$$

and we arrive, through a further application of identity (12), at

$$\frac{1}{2}\delta_{t+}\|p^{n-1}\|^2 = -\frac{1}{2}\delta_{t+}\langle q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle - \frac{B}{4}\delta_{t+}\langle q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle^2$$

from which we can see that the quantity $\mathcal{E}_{KC}^n$, defined by

$$\mathcal{E}_{KC}^n = \mathcal{E}^n + \frac{B}{4}\left(\|q^{n+\frac{1}{2}}\|^2 - \lambda h\langle\delta_{x+}p^n, q^{n+\frac{1}{2}}\rangle\right)^2 \tag{26}$$

satisfies

$$\delta_{t+}\mathcal{E}_{KC}^{n-1} = 0 \qquad \Rightarrow \qquad \mathcal{E}_{KC}^n = \text{constant}$$

Note that $\mathcal{E}^n_{KC}$ is defined here in terms of the conserved quantity $\mathcal{E}^n$ for scheme (17) for the transmission line system.

The expression for $\mathcal{G}^{n-\frac{1}{2}}$ given in (25) might lead one to believe that difference scheme (18) would be implicit. It is, however, possible to arrive at an expression for $\mathcal{G}^{n-\frac{1}{2}}$ purely in terms of the previously calculated state variable $q^{n-\frac{1}{2}}_{i+\frac{1}{2}}$; it is given by

$$\mathcal{G}^{n-\frac{1}{2}} = \frac{1 + B\|q^{n-\frac{1}{2}}\|^2}{1 + \frac{B\lambda^2 h^2}{2}\|\delta_{x-} q^{n-\frac{1}{2}}\|^2} \tag{27}$$

As a result, (18) remains a fully explicit algorithm. It is crucial to note that the "obvious" choice for an approximation to G, namely, $\mathcal{G}^{n-\frac{1}{2}} = 1 + B\|q^{n-\frac{1}{2}}\|^2$ does not lead to a simple energy conservation property.

D. Energy and Numerical Stability

In the previous section, we refrained from calling the conserved quantities $\mathcal{E}^n$ and $\mathcal{E}^n_{KC}$ "energies," because we had yet to show that they indeed have the necessary attribute that an energetic measure must have, namely positivity. As we will see, this positivity property is the key to numerical stability guarantees.

Consider again the conserved quantity $\mathcal{E}^n$ associated with the difference scheme (17) for the transmission line system (3). Clearly, we have, from (14) and (15), that

$$|\langle \delta_{x+} p^n, q^{n+\frac{1}{2}} \rangle| \leq \|\delta_{x+} p^n\| \|q^{n+\frac{1}{2}}\| \leq \frac{2}{h}\|p^n\| \|q^{n+\frac{1}{2}}\|$$

which implies that

$$\begin{aligned} \mathcal{E}^n &\geq \frac{1}{2}\left(\|p^n\|^2 + \|q^{n+\frac{1}{2}}\|^2 - 2\lambda\|q^{n+\frac{1}{2}}\|\|p^n\|\right) \\ &= \frac{1}{2}\left((\|p^n\| - \lambda\|q^{n+\frac{1}{2}}\|)^2 + (1-\lambda^2)\|q^{n+\frac{1}{2}}\|^2\right) = \frac{1}{2}\left((\lambda\|p^n\| - \|q^{n+\frac{1}{2}}\|)^2 + (1-\lambda^2)\|p^n\|^2\right) \end{aligned} \tag{28}$$

¿From either of the equivalent forms of (28), we can then say that, for any choice of p^n_i and $q^{n+\frac{1}{2}}_{i+\frac{1}{2}}$,

$$\mathcal{E}^n \geq 0 \qquad \text{if} \qquad \lambda \leq 1$$

and clearly, for the more strict condition $\lambda < 1$, $\mathcal{E}^n = 0$ only when p^n_i and $q^{n+\frac{1}{2}}_{i+\frac{1}{2}}$ vanish identically. Furthermore, if $\lambda > 1$, there is always a choice of p^n and $q^{n+\frac{1}{2}}$ such that $\mathcal{E}^n < 0$. (Choosing, for instance, $p^n_i = (-1)^i$, $i = 0, \ldots, N-1$ for N even, and $q^{n+\frac{1}{2}}_{i+\frac{1}{2}} = \frac{\lambda h}{2}\delta_{x+} p^n_i$ gives an energy of $\mathcal{E}^n = Nh\left(1-\lambda^2\right)/2$, which is clearly negative for $\lambda > 1$.)

The positivity condition on $\mathcal{E}^n$ allows bounds on the norms of $q^{n+\frac{1}{2}}_{i+\frac{1}{2}}$ and p^n_i, in the same way as for the continuous model problem. Under the condition $\lambda < 1$, familiar as the Courant-Friedrichs-Lewy condition for explicit numerical schemes for hyperbolic systems [5], [6], the two forms of inequality (28) further imply that

$$\|p^n\|, \|q^{n+\frac{1}{2}}\| \leq \sqrt{\frac{2\mathcal{E}^n}{1-\lambda^2}} = \sqrt{\frac{2\mathcal{E}^0}{1-\lambda^2}} \tag{29}$$

This is our guarantee of numerical stability (here in an l^2 norm).

The above analysis can be simply extended to deal with the difference scheme (18) for the Kirchhoff-Carrier system (4) as well. By simple inspection of (26), it should be clear that since $\mathcal{E}^n_{KC} \geq \mathcal{E}^n$, then for $\lambda \leq 1$, $\mathcal{E}^n_{KC}$ will also be non-negative. Furthermore, for $\lambda > 1$, we can again make a choice of p^n and $q^{n+\frac{1}{2}}$ such that $\mathcal{E}^n_{KC}$ becomes negative (for instance, $p^n_i = \epsilon(-1)^i$, $i = 0, \ldots, N-1$ for N even, and $q^{n+\frac{1}{2}}_{i+\frac{1}{2}} = \frac{\lambda h}{2}\delta_{x+} p^n_i$, for $\epsilon < \frac{\sqrt{2(\lambda^2-1)}}{\lambda^2\sqrt{BNh}}$). Thus $\lambda \leq 1$ remains a necessary and sufficient condition for the non-negativity of $\mathcal{E}^n_{KC}$. A bound on the solution size also follows immediately from the fact that $\mathcal{E}^n_{KC} \geq \mathcal{E}^n$ and (29), and thus

$$\|p^n\|, \|q^{n+\frac{1}{2}}\| \leq \sqrt{\frac{2\mathcal{E}^n_{KC}}{1-\lambda^2}} = \sqrt{\frac{2\mathcal{E}^0_{KC}}{1-\lambda^2}} \tag{30}$$

It should be possible to tighten this bound considerably through further analysis.

We note that the above bound on $\|q^{n+\frac{1}{2}}\|$ implies a bound on $\mathcal{G}^{n-\frac{1}{2}}$ as defined in (27), namely,

$$\mathcal{G}^{n-\frac{1}{2}} \leq 1 + B\|q^{n-\frac{1}{2}}\|^2 \leq 1 + \frac{2B\mathcal{E}^0_{KC}}{1-\lambda^2} \tag{31}$$

E. Oscillatory Behavior

One might think that, now that the numerical stability of difference scheme (18) has been shown, it may be employed directly. Unfortunately, under certain conditions, even stable solutions to (18) will be plagued with spurious oscillations. In order to examine this behavior, we may consider spatially sinusoidal solutions of the form

$$p_i^n = \hat{p}^n e^{ji\beta h} \qquad\qquad q_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \hat{q}^{n+\frac{1}{2}} e^{j(i+\frac{1}{2})\beta h}$$

where $j = \sqrt{-1}$ and β is a spatial wavenumber. (As the nonlinearity in (18) is not spatially-dependent, we may expect such analysis to be fruitful.) In this case, as the grid function is periodic, we note that β should be restricted to the discrete set of values $\beta = 2\pi k/(Nh)$, for integer k. For such a component, the difference scheme reduces to

$$\begin{aligned} \hat{p}^n &= \hat{p}^{n-1} + 2j\lambda\mathcal{G}^{n-\frac{1}{2}}\sin(\beta h/2)\hat{q}^{n-\frac{1}{2}} \\ \hat{q}^{n+\frac{1}{2}} &= \hat{q}^{n-\frac{1}{2}} + 2j\lambda\sin(\beta h/2)\hat{p}^n \end{aligned}$$

which can be written in matrix form as

$$\begin{bmatrix} \hat{p}^n \\ \hat{q}^{n+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} 1 & 2j\lambda\mathcal{G}^{n-\frac{1}{2}}\sin(\beta h/2) \\ 2j\lambda\sin(\beta h/2) & 1 - 4\lambda^2\mathcal{G}^{n-\frac{1}{2}}\sin^2(\beta h/2) \end{bmatrix} \begin{bmatrix} \hat{p}^{n-1} \\ \hat{q}^{n-\frac{1}{2}} \end{bmatrix}$$

The eigenvalues $\alpha_\pm^n$ of the update matrix at time step n will be

$$\alpha_\pm^n = \nu^n \pm \sqrt{(\nu^n)^2 - 1}, \quad \text{where} \qquad \nu^n = 1 - 2\lambda^2\mathcal{G}^{n-\frac{1}{2}}\sin^2(\beta h/2)$$

When $-1 \leq \nu^n \leq 1$, or, in other words, if

$$\lambda^2\mathcal{G}^{n-\frac{1}{2}}\sin^2(\beta h/2) \leq 1 \tag{32}$$

the eigenvalues are complex, and of unit magnitude. If, however, $\nu < -1$, then both eigenvalues are real and negative; in particular, one will be of magnitude greater than unity. In this case, we expect *sign-flipping* (i.e., oscillations) to occur, from one time step to the next, accompanied by amplification; such behavior is distinctly unphysical, though we remind the reader that the algorithm remains stable if $\lambda \leq 1$. Notice also that if condition (32) is to be violated, it will be violated first at the Nyquist-limit wavenumber $\beta = \pi/h$. Condition (32) reduces to the condition $\lambda \leq 1$ in the linear case (i.e., when $\mathcal{G}^{n-\frac{1}{2}} = 1$); in other words, in the nonlinear case, the stability and oscillation conditions are distinct, and coalesce as the system approaches linearity.

In order for condition (32) to be satisfied for all β, and at all time steps n, we require

$$\lambda^2\mathcal{G}^{n-\frac{1}{2}} \leq 1 \tag{33}$$

This condition, accompanied by the bound (31) on $\mathcal{G}^{n-\frac{1}{2}}$, gives a condition

$$\lambda \leq \sqrt{1 + 2B\mathcal{E}_{KC}^0 - \sqrt{(1 + 2B\mathcal{E}_{KC}^0)^2 - 1}} \tag{34}$$

for non-oscillatory behavior. We note that the above bound may be unnecessarily strict, due to the looseness in the bounds (31) and (30), though as mentioned previously, it should be possible to improve upon these bounds. We also note that for non-oscillatory behavior, the initial energy of the system intervenes in a way that it did not for the consideration of stability (in which case, we had $\lambda \leq 1$).

Oscillatory behavior is definitely to be avoided in any sound synthesis application, as it will be perceived as high-frequency buzzing. There is, however, another very serious effect in the case of the present algorithm, due to the form of the approximation $\mathcal{G}$ to G; note that from the original second order formulation (2), $c_0\sqrt{G}$ will be a rough (in the frozen-coefficient sense [6]) measure of the speed of wave propagation. This is essentially the point of view taken in digital waveguide models for tension-modulated strings [2], [4]. Our approximation $\mathcal{G}$, however, from (27), depends not just on q, but also on its approximate spatial derivative $\delta_{x-}q$; if oscillatory behavior is present at the spatial Nyquist limit, then this derivative will be large, and the approximation will be unnaturally low, resulting in a reduced wave speed. We examine such aberrant behavior in a numerical example in Section IV.

Some other general comments are in order here regarding these spurious oscillations, which present us with some intriguing open questions. Such oscillatory behavior is a feature of centered difference methods when applied to lossless model problems, even in the linear case [5], [6]; it is particularly problematic when non-smooth initial data are presented to such an algorithm, or in nonlinear generalizations in which discontinuities develop (as in, say, fluid dynamics problems involving shock formation). The reason for this, in the linear case, is that for centered leap-frog type schemes, there is

no damping of any spatial frequency components. Thus, when a discontinuity (made up of contributions from the entire range of spatial frequencies) is present, it will undergo dispersion (when using step sizes away from the Courant limit), and a Gibbs-type phenomenon results, with ringing occurring, particularly in the region containing the discontinuity. We would like to argue, here, that the oscillations that can occur in this algorithm for the Kirchhoff-Carrier equations are of a different nature. There are several indications of this. First, parasitic oscillations occur even for smooth initial data, and occur uniformly across the entire spatial domain of the problem. Second, in this nonlinear case, the amplitude of the parasitic oscillations can actually grow with time; though this may seem contrary to the idea of a numerically stable energy-conserving scheme, it is not: energy growth at high spatial wavenumbers must be accompanied by energy loss at the lower wavenumbers. As a consequence, the amplification of the oscillations is not unbounded. Third, we note that in the linear case, one solution to the problem of spurious oscillation involves the use of so-called artificial viscosity [6], effectively damping high frequencies in a way consistent with the model problem (often, but not always sacrificing accuracy in the process). We have attempted to implement a standard artificial viscosity technique for the nonlinear scheme presented here, but it does not appear to have any effect on spurious oscillations.

Finally, we note that the Kirchhoff-Carrier equation is not, strictly speaking, a hyperbolic equation; due to the spatial averaging of the string tension, the wave speed at any point is dependent on the entire string length, and there is thus not a finite cone of dependence. One reasonable guess as to the nature of the spurious oscillations is then that they are related to oscillations in numerical schemes for other systems of mixed hyperbolic/parabolic character, such as the advection-diffusion system. In such cases, there is an additional condition on the grid spacing h (a cell Peclet condition [5]) in order to avoid oscillatory behavior, independently of stability considerations. Yet in this case, the extra restriction is on λ, not h. We do not, at this time, have a complete understanding of these issues, which clearly merit further attention.

F. Fixed Boundary Conditions

In all the above analysis, we have assumed periodic boundary conditions; here we extend the analysis to cover simple fixed terminations described by

$$p(0,t) = p(L,t) = 0 \tag{35}$$

These are simply incorporated into difference scheme (17) for the transmission line system through a slight modification to (17a); it holds as written for $i = 1, \ldots, N-1$, but at $i = 0$ is specialized to $p_0^n = 0$; the periodicity condition then takes care of the other boundary condition (i.e., $p_N = 0$). (17b) then remains as before, for all $i = 0, \ldots, N-1$; in particular, the spatial difference operator may still be interpreted as operating periodically on p_i^n, given the extra restriction on p_0^n.

(19) and (20) hold as before for $i = 1, \ldots, N-1$, and also hold trivially for $i = 0$, due to the restriction that $p_0^n = 0$. The rest of the energetic analysis is unchanged as well, and this extends to the case of the difference scheme (18) for the Kirchhoff-Carrier system. $\mathcal{G}^{n-\frac{1}{2}}$ may be chosen, as before, by (25) in order to ensure perfect energy conservation, where the energy function is again (26). The stability analysis also follows as before. One slight difference is that for fixed boundary conditions, the form of $\mathcal{G}^{n-\frac{1}{2}}$ to be used in the explicit update will no longer be (27), but must be altered to

$$\mathcal{G}^{n-\frac{1}{2}} = \frac{1 + B\|q^{n-\frac{1}{2}}\|^2}{1 + \frac{\lambda^2 h^2 B}{2} \sum_{i=1}^{N-1} h(\delta_{x-} q_{i+\frac{1}{2}}^{n-\frac{1}{2}})^2} \tag{36}$$

Though boundary conditions are not the focus of this article, we would like to note that the modelling of realistic boundary conditions through standard techniques is by no means simple; frequency domain techniques such as GKS type theories [31], [32], [5] are difficult to apply. Fixed boundary conditions such as those discussed above do provide a useful first approximation to a terminated string under tension (and free boundary conditions, with $q = 0$ at the boundaries are equally simply dealt with). But a more realistic model will model boundary conditions of the so-called lumped impedance type, and it is of critical importance that a string boundary condition be formulated in an energy conserving way, much as in the linear analysis discussed in [16]. A correct formulation of the continuous problem will then necessarily yield a stored energy function containing both the energy of the string and the boundary termination. Under lossless conditions (i.e., if the impedance is purely reactive, without any loss) then this combined energy function will be conserved. It is then necessary to perform discretization of the boundary condition such that a discrete energy function is conserved as well. Though we have not explored such boundary conditions in detail, we do mention that a useful point of departure may be through an application of wave digital filtering principles [33] to the lumped terminations; this has been done in the context of lumped excitation mechanisms for musical instruments in order to yield perfect energy conserving numerical methods based on digital waveguides in [34], [35]. Given the various equivalences between digital waveguides, wave digital filters, and finite difference schemes [36], this would appear to be a worthy direction of future research.

IV. Numerical Experiments

We provide here a few simulation results using difference scheme (18), using the choice (36) as an approximation to G. First, in Figure 1, we show results for a string which is plucked at its center (i.e., initial displacement of a triangular shape, and initial velocity zero), for different plucked amplitudes of 1 cm, 2 cm, 5 cm and 10 cm, and for material parameters as given in the caption[3]. At an amplitude of 1 cm, the behavior of the string is essentially linear, but at higher amplitudes, nonlinear effects quickly become noticeable, in particular, the increased speed of propagation of the wave front, as well as a distortion of the trapezoidal shape which occurs in the linear case. In Figure 2, we show, for the same string and for the same set of plucked initial conditions, time waveforms taken from the displacement at the string center; as expected, for low amplitudes, the waveform is essentially triangular, as in the linear case, but for higher amplitudes becomes distorted (and exhibits a higher frequency of oscillation). Also shown in Figure 2 are plots of the difference between the initial energy $\mathcal{E}_{KC}^0$ contained by the difference scheme and subsequent values $\mathcal{E}_{KC}^n$; notice that in all cases any variation is in the last bits of the machine floating point number representation.

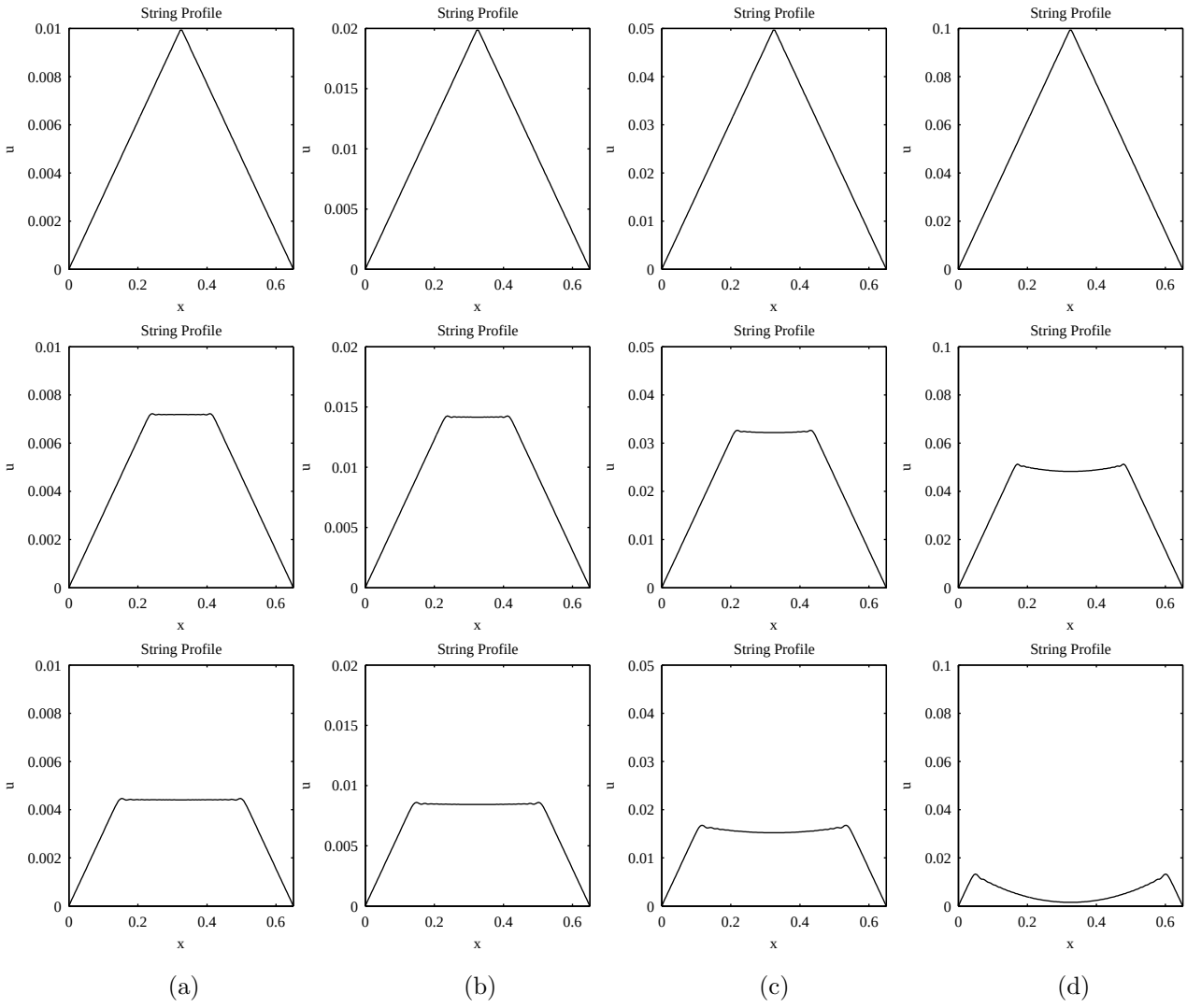

Fig. 1. Time evolution of the profile of a string described by a Kirchhoff-Carrier equation, under the application of the energy-conserving difference scheme (18) accompanied by (36); the string is of length 0.65 m, made of steel (of linear density $\rho = 6 \times 10^{-4}$ kg/m and with Young's Modulus $E = 2 \times 10^{11}\text{N/m}^2$), of cross-sectional area $A = 3.6 \times 10^{-8}\text{m}^2$, under tension $T_0 = 120\text{N}$, and is subject to a triangular (center-plucked) initial condition. Snapshots are taken at times $t = 0$, $t = 0.0002$ s and $t = 0.0004$ s, for a variety of initial amplitudes: (a) 0.01 m, (b) 0.02 m, (c) 0.05 m and (d) 0.1 m. Values of $\lambda = 0.5$, and a sample rate of 200 kHz are used in all cases.

In Figure 3, we explore the effect of variation of λ, with other parameters held constant (the excitation is described in the caption to the figure, and the string parameters are as given in the caption to Figure 1). Recall, from (33), that the

[3]Thanks to Cumhur Erkut at the Helsinki University of Technology for providing these values.

scheme has the potential to develop oscillations at the spatial Nyquist-limit frequency, for values of $\lambda > 1/\sqrt{\max(\mathcal{G})}$, even though the energy function $\mathcal{E}_{KC}$ remains positive and conserved. In the case of the initial condition shown in Figure 3, the maximum value of $\mathcal{G}$ turns out to be approximately 3, implying that we should choose $\lambda \leq 0.58$ in order to avoid such oscillations. We have shown simulations for values of $\lambda = 0.45, 0.55, 0.8$ and 0.95, and note that oscillations begin to appear in the latter two cases. As mentioned at the end of Section III-E, one side-effect of the production of these oscillations will be a reduction in the overall speed of propagation, which is shown clearly for these latter two cases by the resulting lag of the propagating wavefront with respect to the solutions given for $\lambda = 0.45$ and 0.55. We also note that the upper bound given by (34) for the prevention of oscillation, will, in this case, be $\lambda = 0.30$, which is overly conservative.

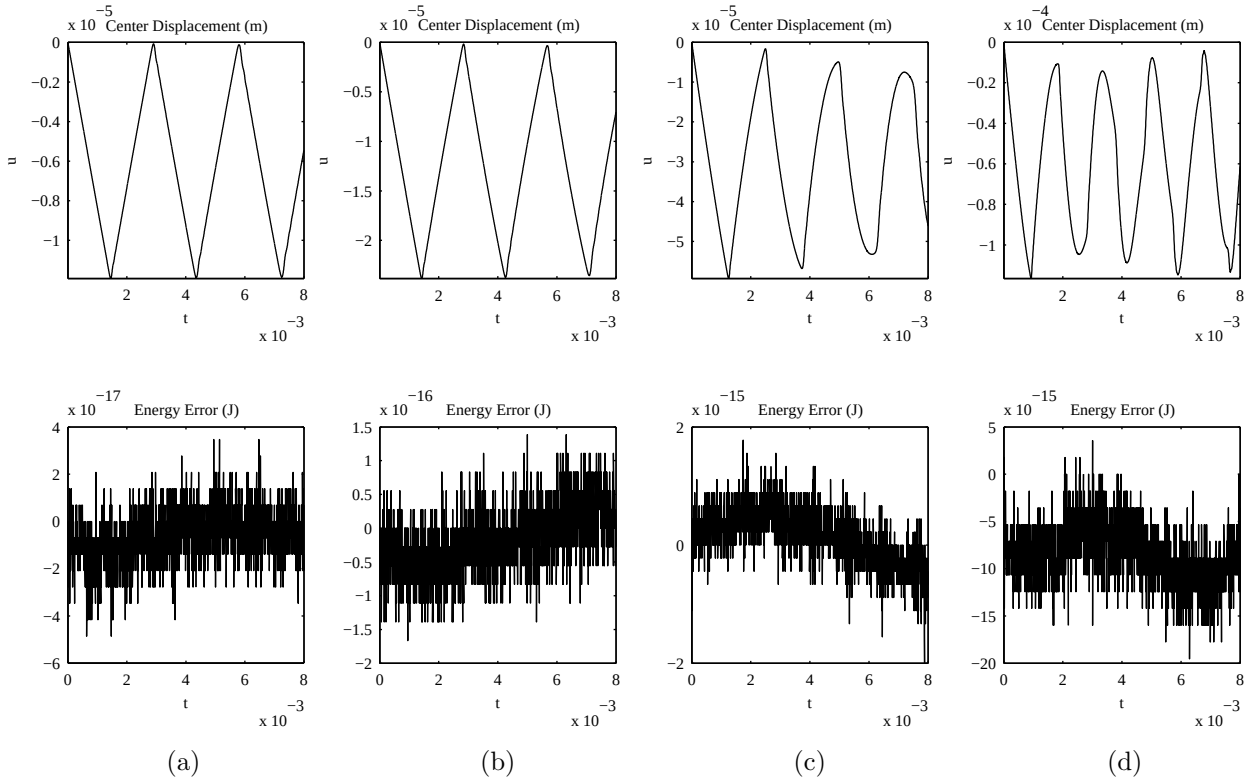

Fig. 2. Time waveforms and energy error, for the string of parameters as given in the caption to Figure 1, for triangular (center-plucked) initial conditions, of amplitudes (a) 0.01 m, (b) 0.02 m, (c), 0.05 m, and (d) 0.1 m. The displacement of the center of the string is plotted as a function of time, over the interval $t \in [0, 0.008]$s. In each case, the difference between the current energy $\mathcal{E}_{KC}$ and the initial value of this energy is plotted as a function of time over the same interval. We have again taken a value of $\lambda = 0.5$ and a sample rate of 200 kHz in all cases.

V. Conclusions

We have shown how perfectly energy-conserving difference schemes for a nonlinear string equation may be derived, and how this property leads to a numerical stability condition. Such robustness of a difference scheme is extremely useful in the context of musical sound synthesis and other applications in which the algorithm must be able to handle a large variety of possible excitations; here, stability depends only on the Courant number, not on any initial conditions applied. This energy conservation property follows directly from energy conservation in the model equation itself, and depends on a particular choice of difference model for the nonlinearity (namely, the approximation $\mathcal{G}$ to G). The analysis we have presented does not depend in any way on frequency domain considerations, though as remarked earlier, due to the spatial invariance of the nonlinearity in question, we could well have introduced spatial Fourier transforms, and performed essentially the same analysis in the transform domain. The nonlinearity persists, however, and a full von Neumann type analysis is impossible in this case. (It is interesting to note that in this special case, spatial Fourier transforms lead naturally to a spectral or pseudospectral type treatment of the system [37], [38]; this direction has been explored, in the case of the Kirchhoff-Carrier equation, in [39].) For more complete nonlinear string models, however, any type of spectral analysis is unrevealing, so it would seem to be best to stick to a pure time/space energy-based

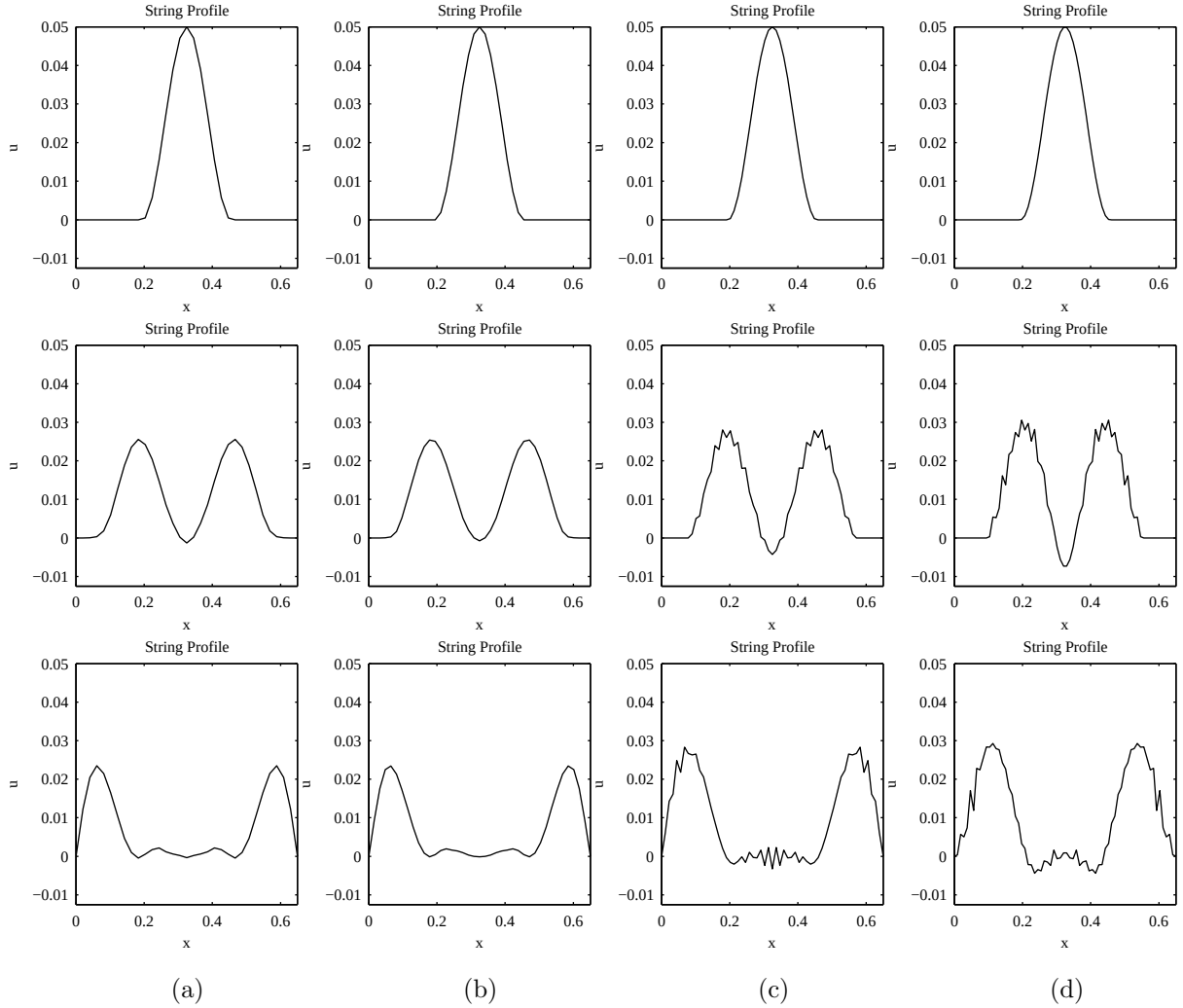

Fig. 3. Time evolution of string profiles, for the string defined by parameters given in the caption to Figure 1, according to difference scheme (18) accompanied by (36), for varying values of λ, and for a sample rate of 50 kHz. In each case, the excitation is an initial string shape of the form of a raised cosine, amplitude 0.05 m, width 0.26 m, at the string center, and snapshots are taken at $t = 0$ s, $t = 0.0002$ s and $t = 0.0004$ s. The different values of λ are (a) 0.45, (b) 0.55, (c) 0.8 and (d) 0.95.

analysis.

We again remark that extended digital waveguide models [2], [3], [1] have been proposed as a means of simulating a tension-modulated string; here, the solution is modelled in terms of travelling waves of variable speed (leading to variable-length digital delay lines). It is important to note that though the Kirchoff-Carrier equation does possess unidirectional travelling wave solutions, the solution can not be modelled as a sum of two travelling waves travelling in opposite directions; distortion of wavefronts will inevitably occur when waves meet. In this sense, then, the extended digital waveguide model must be considered as a rough approximation this system. On the other hand, the Kirchhoff-Carrier Equation itself is an approximation; there are obviously many issues which need clarification, not least of which is the relative perceptual significances of such approximations in sound synthesis applications.

Appendix

I. Introducing Loss

The Kirchhoff-Carrier equation (2) is lossless, at least with respect to the energy definition (10). Adding a simple linear loss term gives

$$\rho \frac{\partial^2 u}{\partial t^2} = \left(T_0 + \frac{EA}{2L} \int_0^L \left(\frac{\partial u}{\partial x} \right)^2 dx \right) \frac{\partial^2 u}{\partial x^2} - \sigma' \frac{\partial u}{\partial t} \tag{37}$$

where $\sigma' \geq 0$ is a constant. The resulting first order system is then

$$\frac{\partial p}{\partial t} = c_0 G \frac{\partial q}{\partial x} - \sigma p \tag{38a}$$

$$\frac{\partial q}{\partial t} = c_0 \frac{\partial p}{\partial x} \tag{38b}$$

where G again is defined as in (4c), and where $\sigma = \sigma'/\rho$. Proceeding through the same energetic analysis as before, we arrive at

$$\frac{dE_{KC}}{dt} = -\sigma \|p\|^2$$

where E_{KC} is again defined as in (10). In other words, for any time $t \geq 0$, we must have

$$E_{KC}(t) \leq E_{KC}(0)$$

and the energy is thus non-increasing.

One choice of interleaved difference scheme involves a so-called semi-implicit approximation [5] to the loss term in (38), i.e.,

$$\delta_{t+} p_i^{n-1} = c_0 \mathcal{G}^{n-\frac{1}{2}} \delta_{x-} q_{i+\frac{1}{2}}^{n-\frac{1}{2}} - \sigma \mu_{t+} p_i^{n-1} \tag{39a}$$

$$\delta_{t+} q_{i+\frac{1}{2}}^{n-\frac{1}{2}} = c_0 \delta_{x+} p_i^n \tag{39b}$$

Again multiplying (39a) by $h\mu_{t+} p_i^n$ and summing over i gives, after several steps,

$$\delta_{t+} \mathcal{E}_{KC}^{n-1} = -\sigma \|\mu_{t+} p^{n-1}\|^2$$

where we have made the same choice of $\mathcal{G}^{n-\frac{1}{2}}$ as (25) and $\mathcal{E}_{KC}^n$ is again defined as in (26). This then implies that

$$\delta_{t+} \mathcal{E}_{KC}^{n-1} \leq 0 \qquad \Rightarrow \qquad \mathcal{E}_{KC}^n \leq \mathcal{E}_{KC}^{n-1} \leq \mathcal{E}_{KC}^0$$

and thus the energy function is non-increasing in the difference realization as well.

The alternate form of $\mathcal{G}^{n-\frac{1}{2}}$ given in (27) must be modified, in the case of loss, to

$$\mathcal{G}^{n-\frac{1}{2}} = \frac{1 + B\|q^{n-\frac{1}{2}}\|^2 - \frac{B\lambda h k \sigma}{2+k\sigma} \langle \delta_{x+} p^{n-1}, q^{n-\frac{1}{2}} \rangle}{1 + \frac{B\lambda^2 h^2}{2+k\sigma} \|\delta_{x-} q^{n-\frac{1}{2}}\|^2}$$

II. Implicit Schemes

It is possible to extend the energetic analysis of so-called explicit schemes (such as (17) and (18)) to the implicit case as well. Implicit numerical methods, though more computationally demanding, may provide certain benefits, especially in terms of the reduction of numerical dispersion [11], [40][4]. In addition, it may be possible to improve on the rather strict conditions for non-oscillatory behavior presented in Section III-E. We present here an implicit scheme for the Kirchhoff-Carrier system (4):

$$\delta_{t+} p_i^{n-1} = c_0 \mathcal{G}^{n-\frac{1}{2}} \delta_{x-} \tilde{q}_{i+\frac{1}{2}}^{n-\frac{1}{2}} \tag{40a}$$

$$\delta_{t+} q_{i+\frac{1}{2}}^{n-\frac{1}{2}} = c_0 \delta_{x+} p_i^n \tag{40b}$$

where $\tilde{q}_{i+\frac{1}{2}}^{n-\frac{1}{2}}$ is defined by

$$\tilde{q}_{i+\frac{1}{2}}^{n-\frac{1}{2}} = \alpha q_{i+\frac{1}{2}}^{n-\frac{1}{2}} + \frac{1-\alpha}{2} \left(q_{i+\frac{1}{2}}^{n+\frac{1}{2}} + q_{i+\frac{1}{2}}^{n-\frac{3}{2}} \right)$$

This scheme depends on a free parameter α and simplifies to the explicit scheme (18) when $\alpha = 1$.

[4] In the linear case, the explicit scheme for the wave equation system yields an exact solution for $\lambda = 1$, i.e., when we are at the Courant limit. Away from this bound, the scheme will introduce significant numerical dispersion. Though it is difficult to clearly define numerical dispersion in the nonlinear case, we note that for a nonlinear system such as the Kirchhoff-Carrier system, the wave speed (if it indeed can be defined) is time-dependent, yet the space step time step ratio remains fixed. One might then argue, heuristically, that it is useful to have a method which behaves well over a large range of wavenumbers.

The energetic analysis follows in a similar fashion to that of the explicit scheme. Multiplying (40a) by $h\mu_{t+}p_i^{n-1}$, and summing over i gives

$$\frac{1}{2}\delta_{t+}\|p^{n-1}\|^2 = c_0\mathcal{G}^{n-\frac{1}{2}}\langle\mu_{t+}p^{n-1}, \delta_{x-}\tilde{q}_{i+\frac{1}{2}}^{n-\frac{1}{2}}\rangle$$

and, after proceeding through the same set of steps as in the explicit case, we arrive at

$$\begin{aligned}\frac{1}{2}\delta_{t+}\|p^{n-1}\|^2 &= -\frac{\mathcal{G}^{n-\frac{1}{2}}}{2k}\langle q^{n+\frac{1}{2}} - q^{n-\frac{3}{2}}, \tilde{q}_{i+\frac{1}{2}}^{n-\frac{1}{2}}\rangle \\ &= -\frac{\mathcal{G}^{n-\frac{1}{2}}}{2}\delta_{t+}\left(\alpha\langle q^{n-\frac{1}{2}}, q^{n-\frac{3}{2}}\rangle + \frac{1-\alpha}{2}\left(\|q^{n-\frac{3}{2}}\|^2 + \|q^{n-\frac{1}{2}}\|^2\right)\right)\end{aligned}$$

Making the choice of

$$\mathcal{G}^{n-\frac{1}{2}} = 1 + B\mu_{t+}\left(\alpha\langle q^{n-\frac{3}{2}}, q^{n-\frac{1}{2}}\rangle + \frac{1-\alpha}{2}\left(\|q^{n-\frac{3}{2}}\|^2 + \|q^{n-\frac{1}{2}}\|^2\right)\right)$$

leads to a conserved quantity given by

$$\begin{aligned}\mathcal{E}_{KC}^n = \quad & \frac{1}{2}\|p^n\|^2 + \frac{\alpha}{2}\langle q^{n+\frac{1}{2}}, q^{n-\frac{1}{2}}\rangle + \frac{1-\alpha}{4}\left(\|q^{n+\frac{1}{2}}\|^2 + \|q^{n-\frac{1}{2}}\|^2\right) \\ + \quad & \frac{B}{4}\left(\alpha\langle q^{n+\frac{1}{2}}, q^{n-\frac{1}{2}}\rangle + \frac{1-\alpha}{2}\left(\|q^{n+\frac{1}{2}}\|^2 + \|q^{n-\frac{1}{2}}\|^2\right)\right)^2\end{aligned}$$

In order to find the conditions on λ under which this quantity is positive, it is useful to introduce the grid function r defined by

$$r = q^{n+\frac{1}{2}} - \frac{\lambda h}{2}\delta_{x+}p^n$$

in which case $\mathcal{E}_{KC}^n$ can be written, after some algebra, as

$$\mathcal{E}_{KC}^n = \frac{1}{2}\left(\|p^n\|^2 + \|r\|^2 + \frac{1-2\alpha}{4}\lambda^2h^2\|\delta_{x+}p^n\|^2\right) + \frac{B}{4}\left(\|r\|^2 + \frac{1-2\alpha}{4}\lambda^2h^2\|\delta_{x+}p^n\|^2\right)^2$$

which is a function purely of the current state variables p^n and $q^{n+\frac{1}{2}}$ (through r). Clearly, for $\alpha \leq \frac{1}{2}$, $\mathcal{E}_{KC}^n$ is non-negative for any value of λ. For $\alpha > \frac{1}{2}$, a sufficient condition can easily be shown to be $\lambda \leq 1/\sqrt{2\alpha-1}$. That this condition is necessary as well can be shown through the counterexample $r = 0$ and $p_i^n = \epsilon(-1)^i$, for some $\epsilon^2 < \frac{2(-1+(2\alpha-1)\lambda^2)}{BNh(2\alpha-1)^2\lambda^4}$. Under these conditions, it is again possible to find bounds on $\|p^n\|$ and $\|q^{n+\frac{1}{2}}\|$ in terms of $\mathcal{E}_{KC}^n$, which remains constant.

It is interesting, and perhaps to be expected, that the conditions under which spurious oscillations occur is distinct in the case of the implicit scheme. Following an analysis similar to that of Section III-E, we may determine a condition on λ, in terms of wavenumbers β and $\mathcal{G}^{n-\frac{1}{2}}$, namely

$$\mathcal{G}^{n-\frac{1}{2}}\lambda^2\sin^2(\beta h/2)\frac{1-(1-\alpha)^2\mathcal{G}^{n-\frac{1}{2}}\lambda^2\sin^2(\beta h/2)}{1+2(1-\alpha)\mathcal{G}^{n-\frac{1}{2}}\lambda^2\sin^2(\beta h/2)} \leq 1$$

for which the scheme will be non-oscillatory. This reduces to (33) for the explicit scheme, when $\alpha = 1$. For other values of α, the analysis is rather complex, and we do not present it in its entirety here, but do note that the bound can be significantly improved. For the value $\alpha = 1/2$, for instance, the above condition reduces to

$$\lambda^2\mathcal{G}^{n-\frac{1}{2}} \leq 4$$

which is to be compared with (33) for the explicit scheme.

The major problem with the use of an implicit scheme is that the state variables are nonlinearly coupled; updating will presumably have to be done through the use of iterative techniques (as it is often done in the linear case)[5], and also for nonlinear schemes [15]. We leave the full development of such implicit methods to a future work.

References

[1] J. O. Smith III, "Music applications of digital waveguides," Tech. Rep. STAN-M-39, Center for Computer Research in Music and Acoustics (CCRMA), Department of Music, Stanford University, 1987, partly incorporated in to `http://ccrma.stanford.edu/~jos/pasp/`.

[2] T. Tolonen, V. Välimäki, and M. Karjalainen, "Modeling of tension modulation nonlinearity in plucked strings," *IEEE Transactions in Speech and Audio Processing*, vol. 8, pp. 300–310, May 2000.

[3] C. Erkut, *Aspects in Analysis and Model-Based Sound Synthesis of Plucked String Instruments*, Ph.D. thesis, Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, 2002.

[4] V. Välimäki, T. Tolonen, and M. Karjalainen, "Plucked-string synthesis algorithms with tension modulation nonlinearity," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, Arizona, USA, March 1999, vol. 2, pp. 977–980.

[5] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, Calif., 1989.

[6] B. Gustaffson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods*, John Wiley and Sons, New York, 1995.

[7] R. Vichnevetsky and J. Bowles, *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*, SIAM, Philadelphia, 1982.

[8] P. Ruiz, "A technique for simulating the vibrations of strings with a digital computer," M.S. thesis, University of Illinois, 1969.

[9] L. Hiller and P. Ruiz, "Synthesizing musical sounds by solving the wave equation for vibrating objects," *Journal of the Audio Engineering Society*, vol. 19, pp. 462–72, 542–51, 1971.

[10] R. Bacon and J. Bowsher, "A discrete model of a struck string," *Acustica*, vol. 41, pp. 21–7, 1978.

[11] A. Chaigne, "On the use of finite differences for musical synthesis. Application to plucked stringed instruments," *J. d'Acoust.*, vol. 5, no. 2, pp. 181–211, 1992.

[12] A. Chaigne and A. Askenfelt, "Numerical simulations of struck strings. I. A physical model for a struck string using finite difference methods," *Journal of the Acoustical Society of America*, vol. 95, no. 2, pp. 1112–8, Feb. 1994.

[13] J. Bensa, S. Bilbao, R. Kronland-Martinet, and J. O. Smith III, "The simulation of piano string vibration: From physical models to finite difference schemes and digital waveguides," *Journal of the Acoustical Society of America*, vol. 114, no. 2, pp. 1095–1107, 2003.

[14] Richtmyer R and K. Morton, *Difference Methods for Initial Value Problems*, John Wiley and Sons, New York, 1967.

[15] S. Li and L. Vu-Quoc, "Finite difference calculus invariant structure of a class of algorithms for the nonlinear Klein Gordon equation," *SIAM Journal of Numerical Analysis*, vol. 32, pp. 1839–1875, 1995.

[16] L. Rhaouti, A. Chaigne, and P. Joly, "Time-domain modeling and numerical simulation of a kettledrum," *Journal of the Acoustical Society of America*, vol. 105, no. 6, pp. 3545–3562, 1999.

[17] J. D. Markel and A. H. Gray, *Linear Prediction of Speech*, Springer Verlag, New York, 1976.

[18] S. Bilbao, "Energy-conserving finite difference schemes for tension-modulated strings," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, Canada, May 2004.

[19] G. Kirchhoff, *Vorlesungen über Mechanik*, Tauber, Leipzig, 1883.

[20] G. F. Carrier, "On the nonlinear vibration problem of the elastic string," *Quarterly of Applied Mathematics*, vol. 3, pp. 157–165, 1945.

[21] C. L. Frota, "Non-local solutions of a nonlinear hyperbolic partial differential equation," *Portugaliae Mathematica*, vol. 51, no. 3, pp. 455–473, 1994.

[22] I. Liu and M. Rincon, "Effect of moving boundaries on the vibrating elastic string," *Applied Numerical Mathematics*, vol. 47, pp. 159–172, 2003.

[23] J. Limaco Ferrel and L. Medeiros, "Kirchhoff-Carrier elastic strings in noncylindrical domains," *Portugaliae Mthematica*, vol. 56, no. 4, pp. 465–500, 1999.

[24] M. Clark and O. Lima, "On a mixed problem for a coupled nonlinear system," *Electronic Journal of Differential Equations*, vol. 1997, no. 06, pp. 1–11, 1997.

[25] P. Morse and U. Ingard, *Theoretical Acoustics*, Princeton University Press, Princeton, New Jersey, 1968.

[26] C. Vallette, "The mechanics of vibrating strings," in *Mechanics of Musical Instruments*, A. Hirschberg, J. Kergomard, and G. Weinreich, Eds., pp. 116–183. Springer, New York, 1995.

[27] G. Anand, "Large-amplitude damped free vibration of a stretched string," *Journal of the Acoustical Society of America*, vol. 45, no. 5, pp. 1089–1096, 1968.

[28] D. Cheng, *Field and Wave Electromagnetics*, p. 438, Addison-Wesley, second edition, 1990.

[29] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," vol. 14, pp. 302–7, 1966.

[30] A. Taflove, *Computational Electrodynamics*, Artech House, Boston, Massachusetts, 1995.

[31] S. Osher, "Stability of difference approximations of dissipative type for mixed initial boundary value problems. I.," *Mathematics of Computation*, vol. 23, pp. 335–340, 1969.

[32] H.-O. Kreiss, "Initial boundary value problems for hyperbolic systems," *Communications on Pure and Applied Mathematics*, vol. 23, pp. 277–298, 1970.

[33] A. Fettweis, "Wave digital filters: Theory and practice," *Proceedings of the IEEE*, vol. 74, no. 2, pp. 270–327, February 1986.

[34] S. Bilbao, J. Bensa, and R. Kronland-Martinet, "The wave digital reed: A passive formulation," in *Proceedings of the COST-G6 Digital Audio Effects Conference*, London, England, September 2003, pp. 225–230.

[35] S. Bilbao, J. Bensa, R. Kronland-Martinet, and J. O. Smith III, "The wave digital piano hammer," in *Proceedings of the 144th meeting of the Acoustical Society of America*, Cancun, Mexico, December 2002, CD-ROM.

[36] S. Bilbao, *Wave and Scattering Methods for Numerical Simulation*, John Wiley and Sons, Chichester, UK, 2004.

[37] L. Trefethen, *Spectral Methods in Matlab*, SIAM, Philadelphia, Pennsylvania, USA, 2000.

[38] B. Fornberg, *A Practical Guide to Pseudospectral Methods*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge, England, 1995.

[39] S. Bilbao, "Modal-type finite difference schemes for nonlinear strings with an energy-conservation property," in *Proceedings of the COST-G6 Digital Audio Effects Conference*, Naples, Italy, October 2004.

[40] S. Bilbao, "Parameterized families of finite difference schemes for the wave equation," *Numerical Methods for Partial Differential Equations*, vol. 20, no. 3, pp. 463–480, 2004.

Responses to Reviewers' Comments.

We'd like to thank the reviewers for their very careful reading of the manuscript; we've learned a lot from them.

Reviewer 1:

1. Done

2. The main reason for a comparison with the waveguide model is that the Kirchhoff-Carrier equation was first approached this way, at least in the musical context. There has not been any work, again in the musical context, on any other type of numerical method. It is also a very active research topic (see publications by Erkut, Valimaki et al.). It is also true, as mentioned in the conclusion, that waveguides only solve the KC equation in an approximate sense. Several comments have been added, in the introduction, as requested.

3. Done

4. The choice of periodic boundary conditions is just a standard way (see, e.g., [6]) of avoiding a consideration of boundary conditions when they are not the main issue. One could also consider an infinite length string, but then there are additional considerations having to do with existence of norms and so forth. We can't think of a musical significance of such a periodic condition, unless one wants to listen to a looped string (could be interesting). Since the boundary condition is not a physical one, we'd prefer to keep it separate, to avoid confusing the reader. If the reviewer disagrees, we'll move it over with the main set of equations.

5. Footnote added

6. Done

7. Done

8. Comment added in text; we'd rather not change the actual numbering, because the choice of G is only one possible choice, and not even the most obvious one; any consistent choice will do, though one will not necessarily have an energy conservation guarantee.

9. Fixed

10. Fixed

11. We agree; thanks to the reviewer for pointing this out. Comment removed.

12. Fixed

13. Let us address your points here:

Regarding lossless model problems, and the application of centered difference schemes, it is well-known that spurious oscillations (stable ones) do result (see Strikwerda, [5], page 16, 100, in his discussion of leapfrog centered methods, also Gustaffson [6], page 52); artificial viscosity (see [6], page 44) is one cure, at least in some cases. We actually tried artificial viscosity with the KC FD model discussed here, and it doesn't work! We've added a few comments in the paper on this. This is a complex matter, and one well worth discussing in more detail.

Comment 1: We definitely agree that it is not at all clear that the linear type eigenvalue analysis applies directly here. But it is probably the only reasonable way to approach the nonlinear problem; indeed, linearized eigenvalue analysis is used extensively (sometimes called frozen-coefficient analysis) for nonlinear problems, particularly in fluid dynamics . Even the standard texts note that such an approach is to be used "with caution" for nonlinear problems ([7], page 3), implying to us that the issue is not really resolved. It is, however, fairly revealing in this case; one could argue, roughly, that if the nonlinearity is slowly varying, then such an analysis becomes acceptable. We've added some more comments on this in the text. Given the above, it is clear that it is not strictly acceptable to treat G as a parameter here, but it may be the best that can be done, in terms of explaining the spurious oscillations. Spectral type methods are discussed in a subsequent publication (we include a new eference).

Comment 2: A modal type version of this method has been described in a soon-to-appear conference publication (submitted after the submission of this manuscript); we include a new reference to it in the paper.

Comment 3: This is an extremely interesting question, and we don't have a complete answer for it. Some hints are given in a new paragraph in the relevant section; we feel a little uncomfortable about putting them in the paper, though we have done so.

Comment 4: The connection with the stability condition has been elaborated.

14. Additional remark added.

15. Done, employing the special boundary condition form given by (36).

16. We agree about the quadratic dependence: the energy goes roughly as the square of the average amplitude; but in floating point arithmetic, as the energy increases, we lose precision in the lower bits. The machine error is visible in the quantization of the error in the plots. Perhaps this is better described as round off error, rather than machine zero. We have changed the phrase in the text.

17. The parameters are for a Finnish kantele string, and are provided by researchers at the Helsinki University of Technology; a footnote has been added as an acknowledgement.

18. Both values, $\lambda = 0.58$ from (33), and $\lambda = 0.3$, from (34) are given; we do mention that the latter bound is unnecessarily strict.

19. For a full von Neumann analysis, we require z-transforms in the time domain as well; we cannot do this here in a useful way due to the nonlinearity. The spatial independence of the nonlinearity does allow Fourier transforms in space, but not time. In other words, we cannot solve for roots of an amplification polynomial equation to give stability results.

20. Some extra material has been added in the introduction to call attention to this shortcoming of digital waveguides; as this is an important point (and a rather open-ended one), we would like to mention it in the concluding remarks as well.

21. The main point of this appendix is to show that energy methods can be applied usefully to more than simple explicit difference schemes. We agree that in the linear case, the solution generated by the explicit difference scheme will be exact, *but only at the Courant limit* $\lambda = 1$, i.e., when the space step and the time step are related by a factor of the wave speed. Away from $\lambda = 1$, we do indeed have dispersion, and implicit methods may be useful (indeed, most developments of implicit methods use the simple advection equation, for which an exact explicit scheme exists, as a starting point. In the nonlinear case, however, the wave speed (if it can even be defined) is variable, meaning that for a fixed space step time step ratio, we are often away from an optimal relation between h and k. In this case, an implicit scheme may be better. We have added additional commentary in the beginning of the section. The oscillatory behavior can indeed be better; we provide some initial condensed results to show this.

Implicit schemes for nonlinear systems have appeared frequently in the literature; we have provided an additional reference. As far as a dispersion analysis goes, it is rather difficult to even define numerical dispersion in the nonlinear case (the KC equations themselves are dispersive, as shown by the plots).

Reviewer 2

Comment 1: We feel that there may be some confusion of the use of the term "energy method." We consider it to refer to stability checking machinery which does not rely on any Fourier or Laplace based frequency domain analysis, in either the time domain or the space domain. We quote Gustaffson Kreiss and Oliger [6], page 183: "*A more direct way of proving stability is the energy method...in this method, no transformations are used, and the analysis is carried out directly in physical space.*" The references mentioned by the reviewer, while well known and consulted frequently by the authors, do not appear to make use of these techniques. The articles by Lambourg, Chaigne and Matignon, and Doutaut, Matignon and Chaigne employ exclusively frequency domain techniques. The article by Rhaouti, Chaigne and Joly do make use of energetic ideas, in the finite element context, but these ideas are not extended to the fully discrete case, and are not definitively linked with numerical stability. We do understand, however, that in the linear case energetic techniques give similar results to those obtained by frequency domain analysis, through Parseval's relation, but we do not feel that the energy method is explicitly used in these articles as a stability checking method for finite difference schemes. We thus feel that we can include a passing reference to the final paper mentioned here, but not the first two.

If the reviewer is in disagreement over this point, we are happy to examine more specific points of the articles in question, or perhaps other articles in the area of musical acoustics simulation.

Comment 2: More realistic boundary modelling is indeed problematic for both the energy method, and standard Fourier techniques (indeed, we'd go so far as to say that it's a rather large unsolved problem). We thank the reviewer for pointing this out. We'll included some comments as suggested, on impedance type conditions, calling attention to energy conserving methods used to deal with lumped nonlinearities in musical acoustics (essentially active boundary conditions).

Comment 3: The reviewer is of course correct, but we do think it does serve as a valuable visual aid (in particular, the machine quantization error is visible in the plot). If the reviewer prefers that we condense the figure, we will of course do so.